



On the construction of bivariate exponential distributions with an arbitrary correlation coefficient

Bladt, Mogens; Nielsen, Bo Friis

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Bladt, M., & Nielsen, B. F. (2008). *On the construction of bivariate exponential distributions with an arbitrary correlation coefficient*. Technical University of Denmark, DTU Informatics, Building 321. D T U Compute. Technical Report No. 2008-13

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

On the construction of bivariate exponential distributions with an arbitrary correlation coefficient

Mogens Bladt* and Bo Friis Nielsen †

October 1, 2008

Abstract

In this paper we use a concept of multivariate phase-type distributions to define a class of bivariate exponential distributions. This class has the following three appealing properties. Firstly, we may construct a pair of exponentially distributed random variables with any feasible correlation coefficient (also negative). Secondly, the class satisfies that any linear combination (projection) of the marginal random variables is a phase-type distributions, The latter property is potentially important for the development hypothesis testing in linear models. Thirdly, it is very easy to simulate the exponential random vectors.

Keywords: Matrix-exponential; phase-type distribution; bivariate exponential distribution; arbitrary correlation

AMS subject classification: 62H05; 60E10 .

1 Introduction

This paper deals with the construction of bivariate exponential distributions being arbitrarily correlated. We use a construction based on a class of mul-

*Institute for Applied Mathematics and Systems, National University of Mexico, Mexico City, bladt@stats.iimas.unam.mx

†Informatics and Mathematical Modeling, Technical University of Denmark, Kgs. Lyngby, bfn@imm.dtu.dk

tivariate Phase-type distributions. We represent an exponential distribution as a higher order phase-type distribution and then link (correlate) two such phase-type distributions to obtain a multivariate phase-type distribution. In this way we are able to construct a class of bivariate exponential distributions with any feasible correlation coefficient and which furthermore has the property that any linear combination of the correlated exponential distributions is a phase-type distribution. The latter closure property is appealing both from an applications point of view as well as for statistical considerations where hypothesis concerning linear sub-spaces are well defined within the class of phase-type distributions. Also, the simulation of bivariate exponential distributions becomes an almost trivial matter.

A number of bivariate distributions with exponential marginals have been discussed in the literature. The classical distributions suggested by Kibble (1941) (sometimes referred to as the Moran–Downton distribution) and Marshall & Olkin (1967) are restricted to the case of non-negative correlation. Other distributions like the Farlie–Gumbel–Morgenstern distributions see e.g. Kotz, Balakrishnan & Johnson (2000) p.353 and Raftery’s distribution see e.g. Kotz, Balakrishnan & Johnson (2000) p.377 can have both positive and negative correlation but only in a limited range. These distributions are all characterized by having a joint Laplace transform, which is rational. Still other distributions of more complicated structure have been suggested, in particular Gumbel’s type I and type III bivariate exponentials (Kotz, Balakrishnan & Johnson (2000) pp.350-355) but again with some restrictions on the form of the correlation. For a somewhat contrived exception, see the discussion following the description of Raftery’s distribution pp.378-379 in Kotz, Balakrishnan & Johnson (2000). In section 2 we provide the necessary background on univariate Phase-type distributions. In section 3 we introduce a class of multivariate Phase-type distributions with particular address to cross moments and correlation. In section 4 we finally construct the bivariate exponential distributions and show that the class is sufficiently rich to contain arbitrarily (feasible) correlated exponential variables. To our knowledge this is the first construction that makes it possible to obtain negatively correlated exponential variables in a way that can be easily used for further modeling and analysis. The construction cannot immediately be generalized to higher dimensions. In section 5 we present our conclusions.

2 Phase-type distributions

Let $\{X_t\}_{t \geq 0}$ denote a Markov process on a finite state-space $E = \{1, 2, \dots, d, d+1\}$. We assume that the states $1, 2, \dots, d$ are transient while the state $d+1$ is absorbing. Then $\{X_t\}_{t \geq 0}$ has an intensity matrix (infinitesimal generator) Λ on the form

$$\Lambda = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix},$$

where \mathbf{T} is a $d \times d$ matrix of intensities corresponding to the transient states and \mathbf{t} is a d -dimensional column vector. The matrix \mathbf{T} is often referred to as the sub-intensity matrix and \mathbf{t} as the exit (rate) vector. We assume that the Markov process has an initial distribution $(\boldsymbol{\pi}, 0)$ with $\boldsymbol{\pi} = (\pi_1, \dots, \pi_d)$, represented as a row vector, concentrated on $\{1, 2, \dots, d\}$, i.e. $\pi_i = \mathbb{P}(X_0 = i)$ and $\sum_{i=1}^d \pi_i = 1$. Since the rows of an intensity matrix must sum to 0 we have that $\mathbf{t} = -\mathbf{T}\mathbf{e}$, where \mathbf{e} is the d -dimensional column vector of 1's.

Let $\tau = \inf\{t \geq 0 | X_t = d+1\}$ denote the time of absorption in state $d+1$. Then we say that τ has a phase-type distribution with representation $(\boldsymbol{\pi}, \mathbf{T})$ and write $\tau \sim \text{PH}(\boldsymbol{\pi}, \mathbf{T})$, Neuts (1975).

The density of τ is given by $f_\tau(x) = \boldsymbol{\pi} \exp(\mathbf{T}x)\mathbf{t}$, where $\exp(\cdot)$ denotes the matrix exponential. The distribution function is $F_\tau(x) = 1 - \boldsymbol{\pi} \exp(\mathbf{T}x)\mathbf{e}$ while its n 'th moment is given by $n!(-\mathbf{T})^{-n}\mathbf{e}$. For an overview on phase-type distributions, their properties and use in risk theory we refer to e.g. Bladt (1994).

Phase-type distributions are mathematically tractable in the sense that probabilistic reasoning using the underlying Markov process structure provides a powerful framework for establishing exact, and in many cases even explicit, solutions to complicated models. Any distribution on \mathbb{R}_+ may be approximated arbitrarily close by a phase-type distribution by increasing the number of phases (i.e. the number of transient states in the underlying Markov process). Phase-type distributions are, however, not adequate for modeling heavy tailed phenomena, since their tails are decreasing exponentially fast (light tailed).

3 Correlated phase-type distributions

The construction of correlated phase-type random variables is based on a class of multivariate phase-type distributions originally introduced by Kulka-

rni (1989). Consider a Markov jump process $\{X_t\}_{t \geq 0}$. Let $r_i : E \rightarrow \mathbb{R}_+$, $i = 1, \dots, m$ be m different mappings and define new random variables Y_1, \dots, Y_m by

$$Y_i = \int_0^\tau r_i(X_t) dt \quad .$$

Then Y_i is constructed by adding up rewards earned in the different states up to absorption. The Y_i 's are phase-type distributed which can be seen easily by adjusting the intensities adequately, see Kulkarni (1989) for details.

Now define the $d \times m$ matrix $\mathbf{K} = \{k_{ij}\} = \{r_j(i)\}$ and let $\mathbf{Y} = (Y_1, \dots, Y_m)$. Then (see Bladt & Nielsen (2008))

$$\langle \mathbf{a}, \mathbf{Y} \rangle \sim \text{PH}(\boldsymbol{\pi}, \boldsymbol{\Delta}(\mathbf{K}\mathbf{a})^{-1}\mathbf{T}) \quad ,$$

for all non-zero $\mathbf{a} \in \mathbb{R}_+^n$, where $\boldsymbol{\Delta}(\mathbf{v})$ is the diagonal matrix with vector \mathbf{v} as diagonal and $\langle \cdot, \cdot \rangle$ denotes the usual inner product in \mathbb{R}^n . Thus the Laplace transform of $\langle \mathbf{a}, \mathbf{Y} \rangle$ is given by

$$\begin{aligned} L_{\langle \mathbf{a}, \mathbf{Y} \rangle}(s) &= \boldsymbol{\pi} (s\mathbf{I} - \boldsymbol{\Delta}(\mathbf{K}\mathbf{a})^{-1}\mathbf{T})^{-1} \boldsymbol{\Delta}(\mathbf{K}\mathbf{a})^{-1}(-\mathbf{T}\mathbf{e}) \\ &= \boldsymbol{\pi} (s\boldsymbol{\Delta}(\mathbf{K}\mathbf{a})\mathbf{T}^{-1} - \mathbf{I})^{-1} \mathbf{t} \quad , \end{aligned}$$

where \mathbf{I} denotes the identity matrix of appropriate dimension. Since

$$L_{\langle \mathbf{a}, \mathbf{Y} \rangle}(s) = \mathbb{E}(\exp(-s\langle \mathbf{a}, \mathbf{Y} \rangle)) \quad ,$$

it is readily seen that $L_{\langle \mathbf{a}, \mathbf{Y} \rangle}(1)$ is the joint Laplace transform of \mathbf{Y} at $\mathbf{a} = (a_1, \dots, a_m)$.

Theorem 3.1 *The multidimensional Laplace transform of \mathbf{Y} is given by*

$$L_{\mathbf{Y}}(s_1, \dots, s_m) = \boldsymbol{\pi} (\boldsymbol{\Delta}(\mathbf{K}\mathbf{s})\mathbf{T}^{-1} - \mathbf{I})^{-1} \mathbf{t} \quad ,$$

where $\mathbf{s} = (s_1, \dots, s_m)$.

There are no explicit formulae, in general, for the joint density or distribution function constructed in this way. Transform inversion is a difficult matter in higher dimensions and often not feasible due to the appearance of essential singularities. Kulkarni (1989) provides a neat method for calculating the conditional survival function given $X_0 = i$ through a system of linear partial differential equations, which is a matter of routine to solve for numerically.

Example 3.1 Let $X_1, X_2, X_3 \sim \exp(1)$ be exponentially distributed with intensity 1. Define $Y_1 = X_1 + X_3$ and $Y_2 = X_2 + X_3$. Then $\mathbf{Y} = (Y_1, Y_2)$ has a multivariate phase-type distribution with $\boldsymbol{\pi} = (1 \ 0 \ 0)$,

$$\mathbf{T} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad \mathbf{K} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} .$$

The variables Y_1 and Y_2 are obviously positively correlated.

In a more general context there has been numerous definitions of multivariate exponential and Gamma distributions. The example above is an example of two random variables which are marginally Erlang distributed. Their joint distribution is a multivariate phase-type distribution in the sense of Kulkarni (1989). Most bivariate gamma constructions, as presented by e.g. Kotz, Balakrishnan & Johnson (2000) have positive correlation or only limited negative correlation. An exception is Schmeiser and Lal's Bivariate Gamma distribution, but this construction depends on using the inverse cumulative distribution function, which is not easy to treat neither numerically nor analytically.

We now turn to the construction of two phase-type distributed random variables which may be either positively or negatively correlated. Assume that the multivariate phase-type distribution we aim to construct have marginal distributions $\text{PH}(\boldsymbol{\alpha}, \mathbf{S})$ and $\text{PH}(\boldsymbol{\beta}, \mathbf{T})$. Then consider the sub-intensity matrix \mathbf{U} and reward matrix \mathbf{K} given by

$$\mathbf{U} = \begin{pmatrix} \mathbf{S} & \mathbf{D} \\ \mathbf{0} & \mathbf{T} \end{pmatrix} \quad \text{and} \quad \mathbf{K} = \begin{pmatrix} \mathbf{e} & \mathbf{0} \\ \mathbf{0} & \mathbf{e} \end{pmatrix} . \quad (1)$$

Here the dimension of \mathbf{e} equals the corresponding dimensions of \mathbf{S} and \mathbf{T} respectively. Thus Y_1 is the time the Markov jump process underlying the phase-type distribution with sub-intensity matrix \mathbf{U} spends in states corresponding to \mathbf{S} and Y_2 is the time the process spends in the states corresponding to the states in \mathbf{T} . The matrix \mathbf{D} defines the rates of the transition from the \mathbf{S} -states to the \mathbf{T} -states.

The joint distribution of $\mathbf{Y} = (Y_1, Y_2)$ is readily seen to be given by the density

$$f_{\mathbf{Y}}(y_1, y_2) = \boldsymbol{\alpha} e^{\mathbf{S}_{y_1}} \mathbf{D} e^{\mathbf{T}_{y_2}} \mathbf{t} .$$

The corresponding joint Laplace transform is thus given by

$$L_{\mathbf{Y}}(s_1, s_2) = \boldsymbol{\alpha} (s_1 \mathbf{S} - \mathbf{I})^{-1} \mathbf{D} (s_2 \mathbf{I} - \mathbf{T})^{-1} \mathbf{t} . \quad (2)$$

The marginal distributions can be obtained by this formula as

$$\begin{aligned} f_{Y_2}(y_2) &= \int_0^\infty \pi e^{\mathbf{S} y_1} \mathbf{D} e^{\mathbf{T} y_2} \mathbf{t} dy_1 \\ &= \boldsymbol{\alpha} (-\mathbf{S})^{-1} \mathbf{D} e^{\mathbf{T} y_2} \mathbf{t} . \end{aligned}$$

A sufficient condition for $Y_2 \sim \text{PH}(\boldsymbol{\beta}, \mathbf{T})$ is hence that $\boldsymbol{\beta} = \boldsymbol{\alpha} (-\mathbf{S})^{-1} \mathbf{D}$. Integrating out the other variable results in the condition $\mathbf{D} \mathbf{e} = \mathbf{s}$, which is equivalent to $\mathbf{u} = -\mathbf{U} \mathbf{e}$ being zero on the entries corresponding to the \mathbf{S} indices.

Theorem 3.2 *The joint moments are given by*

$$\mathbb{E}(Y_1^{n_1} Y_2^{n_2}) = n_1! n_2! \boldsymbol{\alpha} (-\mathbf{S})^{-n_1-1} \mathbf{D} (-\mathbf{T})^{-n_2-1} \mathbf{t} . \quad (3)$$

In particular, the covariance between Y_1 and Y_2 is

$$\text{Cov}(Y_1, Y_2) = \boldsymbol{\alpha} (-\mathbf{S})^{-1} ((-\mathbf{S})^{-1} \mathbf{D} - \mathbf{e} \boldsymbol{\beta}) (-\mathbf{T})^{-1} \mathbf{e} .$$

Proof . Follows immediately differentiating (2). In deriving the formula for the covariance we assume that $\boldsymbol{\beta} = \boldsymbol{\alpha} (-\mathbf{S})^{-1} \mathbf{D}$. Q.E.D.

Example 3.2 Let $\boldsymbol{\alpha} = \boldsymbol{\beta} = (1 \ 0 \ 0 \ 0)$ and

$$\mathbf{S} = \mathbf{T} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} .$$

and let \mathbf{U} and \mathbf{K} be defined as in (1). Then the conditions $\boldsymbol{\beta} = -\boldsymbol{\alpha} \mathbf{S}^{-1} \mathbf{D}$ and $\mathbf{D} \mathbf{e} = \mathbf{s}$ implies that the only possible form of \mathbf{D} is the matrix

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} .$$

This corresponds to the marginal distributions being independent. Thus the only possible multivariate phase-type distribution with marginal 4-dimensional Erlang distribution which can be represented by an eight-dimensional \mathbf{U} matrix is the one which consists of two independent 4-dimensional Erlang distributions.

In order to incorporate a non-trivial correlation structure it will be necessary to represent the same distribution in a matrix \mathbf{U} of higher dimension.

4 Construction of bivariate exponential distributions

The example from the previous section also applies to the case of a bivariate exponential distribution, i.e. the only bivariate exponential distribution which can be obtained using a two dimensional \mathbf{U} matrix is the trivial one consisting of two independent exponentially distributed random variables. The following lemmas shall provide a method for introducing correlation between the variables in the construction using the \mathbf{U} matrix.

The following results are well known and easily proven.

Lemma 4.1 *If $X \sim PH(\boldsymbol{\pi}, \mathbf{T})$ and $-\mathbf{T}\mathbf{e} = \lambda\mathbf{e}$ for some constant $\lambda > 0$, then X is exponentially distributed with intensity λ .*

Lemma 4.2 *Let $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_n)$ be any initial vector and*

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & \lambda_1 - \lambda & 0 & \dots & 0 \\ 0 & -\lambda_2 & \lambda_2 - \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & -\lambda \end{pmatrix},$$

where it is assumed that $\lambda < \lambda_i$ for all i . Then $PH(\boldsymbol{\pi}, \mathbf{T})$ is equivalent to an exponential distribution with rate λ .

We shall denote this special form of the matrix \mathbf{T} by $\mathbf{E}(\boldsymbol{\lambda})$, where $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{n-1}, \lambda)$.

Another method for constructing an exponential distribution from a higher order phase-type distribution is by time-reversing the construction above.

We refer to Andersen, Neuts & Nielsen (2004) and Ramaswami (1990) for details on time reversing general phase-type distributions. However, for the case where $\boldsymbol{\pi} = (1 \ 0 \ 0 \ \dots \ 0)$ the time reversed is particularly simple

Lemma 4.3 *Let $\alpha = (\alpha_1, \dots, \alpha_n)$ with*

$$\alpha_i = \frac{\lambda}{\lambda_{n-i+1}} \prod_{j=1}^{n-i} \frac{\lambda_j - \lambda}{\lambda_j}$$

and

$$\mathbf{S} = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & 0 \\ 0 & -\lambda_{n-1} & \lambda_{n-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_1 \end{pmatrix}, \quad (4)$$

where $\lambda < \lambda_i$ for all i . Then $\text{PH}(\boldsymbol{\alpha}, \mathbf{S})$ represents an exponential distribution with rate λ . We shall refer to the matrix (4) as $\mathbf{Er}(\boldsymbol{\lambda})$, where $\boldsymbol{\lambda} = (\lambda, \lambda_{n-1}, \dots, \lambda_1)$.

We apply these Lemmas to the construction of bivariate exponential distributions with an arbitrary correlation coefficient.

Let Y_1 and Y_2 be exponential distributions with intensities λ and μ respectively. For a given $n \in \mathbb{N}$ we may write

$$Y_1 \sim \text{PH}(\boldsymbol{\pi}, \mathbf{Er}(\boldsymbol{\lambda})) \text{ and } Y_2 \sim \text{PH}(\boldsymbol{\alpha}, \mathbf{Er}(\boldsymbol{\mu})), \quad (5)$$

where $\boldsymbol{\lambda} = (n\lambda, (n-1)\lambda, \dots, \lambda)$, $\boldsymbol{\mu} = (\mu, 2\mu, \dots, n\mu)$, $\boldsymbol{\pi} = (1, 0, \dots, 0)$, and $\boldsymbol{\alpha} = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$.

Now we construct the joint density by appropriately specifying the matrix $\mathbf{D} = \lambda \mathbf{P}$, and letting $\mathbf{P} = \{p_{ij}\}$. Then the Laplace transform is given by

$$\begin{aligned} L_{\mathbf{Y}}(s_1, s_2) &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n p_{ij} \prod_{\ell=1}^i \frac{n-\ell+1}{s_1 + n-\ell+1} \prod_{k=j}^n \frac{k}{s_2 + k} \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n p_{ij} \prod_{\ell=n-i+1}^n \frac{\ell}{s_1 + \ell} \prod_{k=j}^n \frac{k}{s_2 + k}. \end{aligned}$$

We shall make use of the following partial fraction expansion.

Lemma 4.4

$$\prod_{k=j}^n \frac{k}{s+k} = \sum_{k=j}^n \frac{n!}{(j-1)!(n-k)!(k-j)!} \frac{(-1)^{k-j}}{s+k} .$$

Proof . The formula holds for $n = j$. Assume that it is valid for $n - 1$. Then

$$\begin{aligned} \prod_{k=j}^{n+1} \frac{k}{s+k} &= \prod_{k=j}^n \frac{k}{s+k} \frac{n+1}{s+n+1} \\ &= \left(\sum_{k=j}^n \frac{n!}{(j-1)!(n-k)!(k-j)!} \frac{(-1)^{k-j}}{s+k} \right) \frac{n+1}{s+n+1} \\ &= \sum_{k=j}^n \frac{n!}{(j-1)!(n-k)!(k-j)!} \frac{(-1)^{k-j}}{s+k} \frac{n+1}{s+n+1} \\ &= \sum_{k=j}^n \frac{(n+1)!(-1)^{k-j}}{(j-1)!(n-k)!(k-j)!(n+1-k)!} \left[\frac{1}{s+k} - \frac{1}{s+n+1} \right] \\ &= \sum_{k=j}^n \frac{(n+1)!(-1)^{k-j}}{(j-1)!(n+1-k)!(k-j)!} \frac{1}{s+k} \\ &\quad + \frac{1}{s+n+1} \frac{(n+1)!}{(j-1)!} \sum_{k=j}^n \frac{(-1)^{k+1-j}}{(n+1-k)!(k-j)!} . \end{aligned}$$

Using induction the latter sum is then seen to be equal to $(-1)^{n+1-j}/(n+1-j)!$.
Q.E.D.

Theorem 4.1 *The joint density for $\mathbf{Y} = (Y_1, Y_2)$ is given by*

$$f(y_1, y_2) = \sum_{\ell=1}^n \sum_{k=1}^n c_{\ell k} \ell \lambda e^{-\ell \lambda y_1} k \mu e^{-k \mu y_2},$$

with

$$c_{\ell k} = \frac{(-1)^{\ell+k-(n+1)}}{n} \binom{n}{\ell} \binom{n}{k} \sum_{i=n+1-\ell}^n \sum_{j=1}^k p_{ij} (-1)^{-i-j} \binom{\ell-1}{n-i} \binom{k-1}{k-j} . \quad (6)$$

Proof .

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n p_{ij} \sum_{\ell=n-i+1}^n \frac{n!}{(n-i)!(n-\ell)!(\ell-(n+1-i))!} \frac{(-1)^{\ell-(n+1-i)}}{s+\ell} \times \\
& \quad \sum_{k=j}^n \frac{n!}{(j-1)!(n-k)!(k-j)!} \frac{(-1)^{k-j}}{t+k} \\
&= \frac{1}{n} \sum_{\ell=1}^n \sum_{k=1}^n \frac{n!n!}{(n-\ell)!(n-k)!} \frac{(-1)^{\ell-(n+1)}}{s+\ell} \frac{(-1)^k}{t+k} \times \\
& \quad \sum_{i=n+1-\ell}^n \sum_{j=1}^k p_{ij} \frac{(-1)^{-i-j}}{(n-i)!(\ell-(n+1-i))!(j-1)!(k-j)!} \\
&= \frac{1}{n} \sum_{\ell=1}^n \sum_{k=1}^n \binom{n}{\ell} \binom{n}{k} \frac{(-1)^{\ell-(n+1)} \ell}{s+\ell} \frac{(-1)^k k}{t+k} \times \\
& \quad \sum_{i=n+1-\ell}^n \sum_{j=1}^k p_{ij} \frac{(-1)^{-i-j} (\ell-1)!(k-1)!}{(n-i)!(\ell-(n+1-i))!(j-1)!(k-j)!} \\
&= \frac{1}{n} \sum_{\ell=1}^n \sum_{k=1}^n \binom{n}{\ell} \binom{n}{k} \frac{(-1)^{\ell-(n+1)} \ell}{s+\ell} \frac{(-1)^k k}{t+k} \times \\
& \quad \sum_{i=n+1-\ell}^n \sum_{j=1}^k p_{ij} (-1)^{-i-j} \binom{\ell-1}{n-i} \binom{k-1}{k-j}
\end{aligned}$$

from which the result follows.

Q.E.D.

Corollary 4.1 *If $p_{ij} = \delta_{i-j}$,*

$$c_{\ell k} = 1\{n+1 \leq \ell+k\} \frac{(-1)^{\ell+k-(n+1)}}{n} \binom{n}{\ell} \binom{n}{k} \binom{\ell+k-2}{n-1} .$$

If $p_{ij} = \delta_{i+j-(n+1)}$,

$$c_{\ell k} = \frac{(-1)^{\ell+k}}{n} \binom{n}{\ell} \binom{n}{k} \binom{\ell+k-2}{\ell-1} .$$

Proof . The inner sum of (6) reduces to

$$\sum_{i=n+1-\ell}^k \binom{\ell-1}{n-i} \binom{k-1}{k-i} \text{ and } (-1)^{n+1} \sum_{j=1}^{\min(\ell,k)} \binom{\ell-1}{j-1} \binom{k-1}{j-1} ,$$

which again can be rewritten into

$$\sum_{h=0}^{\ell+k-n-1} \binom{\ell-1}{h} \binom{k-1}{h+n-\ell} \text{ and } (-1)^{n+1} \sum_{h=0}^{\min(\ell-1, k-1)} \binom{\ell-1}{h} \binom{k-1}{h} ,$$

respectively, from which the result follows.

Q.E.D.

Applying (3) to the current setting yields

$$\mathbb{E}(Y_1 Y_2) = \frac{1}{\lambda \mu} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left(\sum_{\ell=1}^i \frac{1}{n+1-\ell} \right) \left(\sum_{\ell=j}^n \frac{1}{\ell} \right) p_{ij} .$$

The maximal negative correlation between two positive random variables is obtained by an antithetic construction. Thus in the case of the exponential distribution $V_1, V_2 \sim \exp(1)$, we obtain maximum negative correlation by letting $V_1 = -\log(U)$ and $V_2 = -\log(1-U)$, where U is uniformly distributed over $[0, 1]$. The maximum negative correlation coefficient between two exponential distributions $\exp(\lambda)$ and $\exp(\mu)$ respectively is hence $1 - \pi^2/6$ as obtained noting that $\int_0^1 \log(u) \log(1-u) du = 2 - \pi^2/6$, see e.g. Kotz, Balakrishnan & Johnson (2000) p. 379.

Theorem 4.2 *Let $p_{ij} = \delta_{i-j}$ and let $Y_1^{(n)}$ and $Y_2^{(n)}$ denote the exponentially distributed random variables (5) based on the dimension $n \in \mathbb{N}$. Then*

$$\text{corr} \left(Y_1^{(n)}, Y_2^{(n)} \right) = 1 - \sum_{i=1}^n \frac{1}{i^2} \rightarrow 1 - \frac{\pi^2}{6} \text{ as } n \rightarrow \infty ,$$

i.e. the maximum negative correlation.

Proof. Without loss of generality we may assume that $\lambda = \mu = 1$. Since $p_{ij} = \delta_{i-j}$ we get that

$$\mathbb{E} \left(Y_1^{(n)} Y_2^{(n)} \right) = \frac{1}{n} \sum_{i=1}^n \left(\sum_{\ell=1}^i \frac{1}{n+1-\ell} \right) \left(\sum_{\ell=i}^n \frac{1}{\ell} \right) .$$

By interchanging the order of summation we see that

$$\begin{aligned}
\mathbb{E} \left(Y_1^{(n)} Y_2^{(n)} \right) &= \frac{1}{n} \sum_{\ell=1}^n \sum_{k=\ell}^n (k - \ell + 1) \frac{1}{n+1-\ell} \frac{1}{k} \\
&= \frac{1}{n} \sum_{k=1}^n \sum_{\ell=1}^k \frac{1}{k} \frac{k - \ell + 1}{n+1-\ell} \\
&= \frac{1}{n} \sum_{k=1}^n \sum_{\ell=1}^k \frac{1}{k} \sum_{\ell=n+1-k}^n \frac{k + \ell - n}{\ell} \\
&= \frac{1}{n} \sum_{k=1}^n \frac{1}{k} \left(k + \sum_{\ell=n+1-k}^n \frac{k}{\ell} - n \sum_{\ell=n+1-k}^n \frac{1}{\ell} \right) \\
&= 1 + \frac{1}{n} \sum_{k=1}^n \sum_{\ell=n+1-k}^n \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+1-k}^n \frac{1}{\ell} \\
&= 1 + \frac{1}{n} \sum_{k=1}^n \sum_{\ell=1}^n 1\{n \geq \ell \geq n+1-k\} \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+1-k}^n \frac{1}{\ell} \\
&= 1 + \frac{1}{n} \sum_{\ell=1}^n \frac{1}{\ell} \sum_{k=1}^n 1\{n \geq \ell \geq n+1-k\} - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+1-k}^n \frac{1}{\ell} \\
&= 1 + \frac{1}{n} \sum_{\ell=1}^n \frac{1}{\ell} \ell - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+1-k}^n \frac{1}{\ell} \\
&= 2 - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+1-k}^n \frac{1}{\ell} .
\end{aligned}$$

We then consider the difference between the $(n+1)$ 'th and n 'th order of the

last term. This amounts to

$$\begin{aligned}
& \sum_{k=1}^{n+1} \frac{1}{k} \sum_{\ell=n+2-k}^{n+1} \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+1-k}^n \frac{1}{\ell} \\
&= \sum_{k=1}^{n+1} \frac{1}{k} \sum_{\ell=n+2-k}^{n+1} \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+1-k}^{n+1} \frac{1}{\ell} + \sum_{k=1}^n \frac{1}{k} \frac{1}{n+1} \\
&= \sum_{k=1}^{n+1} \frac{1}{k} \sum_{\ell=n+2-k}^{n+1} \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \sum_{\ell=n+2-k}^{n+1} \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \frac{1}{n+1-k} + \sum_{k=1}^n \frac{1}{k} \frac{1}{n+1} \\
&= \frac{1}{n+1} \sum_{\ell=1}^{n+1} \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \frac{1}{n+1-k} + \sum_{k=1}^n \frac{1}{k} \frac{1}{n+1} \\
&= \frac{1}{(n+1)^2} + \frac{2}{n+1} \sum_{\ell=1}^n \frac{1}{\ell} - \sum_{k=1}^n \frac{1}{k} \frac{1}{n+1-k} \\
&= \frac{1}{(n+1)^2} + \sum_{k=1}^n \frac{2(n+1-k) - (n+1)}{(n+1)(n+1-k)k} \\
&= \frac{1}{(n+1)^2} + \frac{1}{n+1} \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{n+1-k} \right) \\
&= \frac{1}{(n+1)^2} .
\end{aligned}$$

Hence the result follows noting that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. Q.E.D.

Please notice that the joint distribution of $(-\log(U), -\log(1-U))$ is not a multivariate phase-type distribution itself.

The maximal positive correlation between two positive random variables is of course 1. We now prove that we can reach this maximum correlation in the limit if we specify \mathbf{P} to be the “diagonal” matrix where the diagonal is the running from lower left to the upper right corner of \mathbf{P} .

Theorem 4.3 *Let $p_{ij} = \delta_{i+j-(n+1)}$ and let $Y_1^{(n)}$ and $Y_2^{(n)}$ denote the exponentially distributed random variables (5) based on the dimension $n \in \mathbb{N}$. Then*

$$\text{corr} \left(Y_1^{(n)}, Y_2^{(n)} \right) = 1 - \frac{1}{n} \sum_{k=1}^n \frac{1}{k} \rightarrow 1 \quad \text{as } n \rightarrow \infty ,$$

i.e. the maximum positive correlation.

Proof .

$$\begin{aligned}
\mathbb{E}\left(Y_1^{(n)}Y_2^{(n)}\right) &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left(\sum_{\ell=1}^i \frac{1}{n+1-\ell} \right) \left(\sum_{k=j}^n \frac{1}{k} \right) \delta_{i+j-(n+1)} \\
&= \frac{1}{n} \sum_{i=1}^n \left(\sum_{\ell=1}^i \frac{1}{n+1-\ell} \right) \left(\sum_{k=n+1-i}^n \frac{1}{k} \right) \\
&= \frac{1}{n} \sum_{i=1}^n \left(\sum_{k=n+1-i}^n \frac{1}{k} \right)^2 \\
&= \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^n \sum_{\ell=1}^n 1_{\{\min(k, \ell) \geq n+1-i\}} \frac{1}{k} \frac{1}{\ell} \\
&= \frac{1}{n} \sum_{k=1}^n \sum_{\ell=1}^n \frac{1}{k} \frac{1}{\ell} \min(k, \ell) \\
&= \frac{1}{n} \sum_{k=1}^n \sum_{\ell=1}^k \frac{1}{k} + \frac{1}{n} \sum_{k=1}^n \sum_{\ell=k}^n \frac{1}{\ell} - \frac{1}{n} \sum_{k=1}^n \frac{1}{k} \\
&= 2 - \frac{1}{n} \sum_{k=1}^n \frac{1}{k} ,
\end{aligned}$$

from which the result follows.

Q.E.D.

Remark 4.1 If $p_{ij} = \frac{1}{n}$ we get independent random variables Y_1 and Y_2 .

Finally we construct a two-dimensional exponential distribution with arbitrary correlation coefficient.

Theorem 4.4 Let $\rho \in (1 - \frac{\pi^2}{6}, 1)$. Then we can construct a two-dimensional exponential vector (Y_1, Y_2) with correlation coefficient ρ in the following way. If $\rho > 0$ we choose $n \in \mathbb{N}$ such that $\rho_{\max} = \mathbf{corr}(Y_1^{(n)}, Y_2^{(n)}) > \rho$ and

$$\mathbf{P} = \frac{\rho}{\rho_{\max}} \{\delta_{i+j-n-1}\}_{i,j} + \frac{1-\rho}{\rho_{\max}} \frac{1}{n} \mathbf{e} \mathbf{e}',$$

and if $\rho < 0$ we choose $n \in \mathbb{N}$ such that $\rho_{\min} = \mathbf{corr}(Y_1^{(n)}, Y_2^{(n)}) < \rho$ and

$$\mathbf{P} = \frac{\rho}{\rho_{\min}} \{\delta_{i-j}\}_{i,j} - \frac{1+\rho}{\rho_{\min}} \frac{1}{n} \mathbf{e} \mathbf{e}' .$$

Proof . Follows from additivity of the first cross moment.

Q.E.D.

5 Conclusions

We have constructed a class of bivariate exponential distributions which may match any feasible correlation. The construction is based on one single underlying Markov jump process which makes it a trivial matter to simulate the bivariate realizations. Any linear combination of the marginal exponential variables is Phase-type distributed with a known representation. We have outlined three different choices of \mathbf{D} corresponding to most negatively correlated, independent and most positively correlated cases. There exists an ample space for further specifications of \mathbf{D} which may result in alternative sub-classes of interest. We have shown, however, that with the three somehow extreme cases we may construct variables with any feasible correlation coefficient by appropriate mixing.

Acknowledgments

The authors are grateful for the support by *Otto Mønstedts foundation*, *Danish research council for technology and production sciences grant no. 247-07-0090* and *Sistema Nacional de Investigadores*, grant 15945.

References

- Andersen, A. T.; Neuts, M. F. & Nielsen, B. F. (2004). “On the Time Reversal of Markovian Arrival Processes”. *Stochastic Models*, 20(3):237–260.
- Bladt, M. (1994). “A review on phase-type distributions and their use in risk theory”. *Astin Bulletin*, 35(1):145–161.
- Bladt, M. & Nielsen, B. F. (2008). “Multivariate matrix-exponential distributions: a characterization”. *submitted*.
- Kibble, W. F. (1941). “A two-variate gamma type distribution”. *Sankhyā*, 5:137–150.
- Kotz, S.; Balakrishnan, N. & Johnson, N. L. (2000). *Continuous Multivariate Distributions*. John Wiley and Sons.

- Kulkarni, V. G. (1989). “A new class of multivariate phase type distributions”. *Operations Research*, 37(1):151–158.
- Marshall, A. W. & Olkin, I. (1967). “A Multivariate Exponential Distribution”. *Journal of the American Statistical Association*, 62(317):30–44.
- Neuts, M. F. (1975). “Probability distributions of phase type”. In *Liber Amicorum Professor Emeritus H. Florin*, pages 173–206, Department of Mathematics, University of Louvain, Belgium.
- Ramaswami, V. (1990). “A Duality Theorem for the Matrix Paradigms in Queueing Theory”. *Commun. Statist. -Stochastic Models*, 6(1):151–161.